University of Zurich[UZH]

Zurich Open Repository and Archive
University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2013

Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis

Beecham, A H ; Patsopoulos, N A ; Xifara, D K ; Jelčić, Ilijas ; Martin, Roland ; Sospedra, Mireia

Abstract: Using the ImmunoChip custom genotyping array, we analyzed 14,498 subjects with multiple sclerosis and 24,091 healthy controls for 161,311 autosomal variants and identified 135 potentially associated regions ($P < 1.0 \times 10(-4)$). In a replication phase, we combined these data with previous genome-wide association study (GWAS) data from an independent 14,802 subjects with multiple sclerosis and 26,703 healthy controls. In these 80,094 individuals of European ancestry, we identified 48 new susceptibility variants ($P < 5.0 \times 10(-8)$), 3 of which we found after conditioning on previously identified variants. Thus, there are now 110 established multiple sclerosis risk variants at 103 discrete loci outside of the major histocompatibility complex. With high-resolution Bayesian fine mapping, we identified five regions where one variant accounted for more than 50% of the posterior probability of association. This study enhances the catalog of multiple sclerosis risk variants and illustrates the value of fine mapping in the resolution of GWAS signals.

DOI: https://doi.org/10.1038/ng.2770

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-85424
Journal Article
Accepted Version

Originally published at:
Beecham, A H; Patsopoulos, N A; Xifara, D K; Jelčić, Ilijas; Martin, Roland; Sospedra, Mireia (2013). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nature Genetics, 45(11):1353-1360.
DOI: https://doi.org/10.1038/ng.2770

Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis

International Multiple Sclerosis Genetics Consortium (IMSGC)*

* A full list of authors and affiliations appears at the end of the paper

Introductory paragraph word count: 148

Main text word count: 1539

Online Methods word count: 1542

Number of tables: 3

Number of figures: 2

Number of supplementary tables: 14

Number of supplementary figures: 100

Introductory Paragraph

Using the ImmunoChip custom genotyping array, we analysed 14,498 multiple sclerosis subjects and 24,091 healthy controls for 161,311 autosomal variants and identified 135 potentially associated regions (p-value < 1.0×10^{-4}). In a replication phase, we combined these data with previous genome-wide association study (GWAS) data from an independent 14,802 multiple sclerosis subjects and 26,703 healthy controls. In these 80,094 individuals of European ancestry we identified 48 new susceptibility variants (p-value < 5.0×10^{-8}); three found after conditioning on previously identified variants. Thus, there are now 110 established multiple sclerosis risk variants in 103 discrete loci outside of the Major Histocompatibility Complex. With high resolution Bayesian fine-mapping, we identified five regions where one variant accounted for more than 50% of the posterior probability of association. This study enhances the catalogue of multiple sclerosis risk variants and illustrates the value of fine-mapping in the resolution of GWAS signals.

Main text

Multiple sclerosis (OMIM 126200) is an inflammatory demyelinating disorder of the central nervous system that is a common cause of chronic neurological disability.[1,2] It has its greatest prevalence amongst individuals of Northern European ancestry[3] and is moderately heritable,[4] with a sibling relative recurrence risk (λ_s) of ~ 6.3.[5] Aside from the early success in demonstrating the important effects exerted by variants in the Human Leukocyte Antigen (HLA) genes from the Major Histocompatibility Complex (MHC),[6] there was little progress in unravelling the genetic architecture underlying multiple sclerosis susceptibility prior to the advent of genome-wide association studies (GWAS). Over the last decade, our Consortium has performed several GWAS and meta-analyses in large cohorts,[7-10] cumulatively identifying more than 50 non-MHC susceptibility alleles. As in other complex diseases, available data suggest that many additional susceptibility alleles remain to be identified.[11]

The striking overlap in the genetic architecture underlying susceptibility to autoimmune diseases[9,10,12,13] prompted the collaborative construction of the "ImmunoChip" (see **Supplementary Note** and **Supplementary Figs. 1** and **2** for details of IMSGC nominated content), an efficient genotyping platform designed to deeply interrogate 184 non-MHC loci with genome-wide significant associations to at least one autoimmune disease and provide lighter coverage of other genomic regions with suggestive evidence of association.[14] Here, we report a large-scale effort that leverages the ImmunoChip to detect association with multiple sclerosis susceptibility and refine these associations via Bayesian fine-mapping.

After stringent quality control (QC), we report genotypes on 28,487 individuals of European ancestry (14,498 multiple sclerosis subjects, 13,989 healthy controls) that are independent of previous GWAS efforts. We supplemented these data with 10,102 independent control subjects provided by the International Inflammatory Bowel Disease Genetics Consortium (IIBDGC)[15] bringing the total to 38,589 individuals (14,498 multiple sclerosis subjects and 24,091 healthy controls). We performed variant level QC, population outlier identification, and subsequent case-control analysis in 11 country-organized strata. To account for within-stratum population stratification we used the first five principal components as covariates in the association analysis. Per stratum odds ratios (OR) and respective standard errors (SE) were then combined with an inverse variance meta-analysis under a fixed effects model. In total we tested 161,311 autosomal variants that passed QC in at least two of the 11 strata (**Online Methods**). A circos plot[16] summarising the results from this discovery phase analysis is shown in **Figure 1**.

We defined an *a priori* discovery threshold of p-value $<1 \times 10^{-4}$ and identified 135 primary statistically independent association signals; 67 in the designated fine-mapping regions and 68 in less densely covered regions selected for deep replication of earlier GWAS. Another 13 secondary and 2 tertiary statistically

independent signals were identified by forward stepwise logistic regression. A total of 48 of the 150 statistically independent association signals (**Supplementary Table 1**) reached a genome-wide significance p-value $<5 \times 10^{-8}$ at the discovery phase alone. Next, we replicated our findings in 14,802 multiple sclerosis subjects and 26,703 healthy controls with available GWAS data imputed to the 1000 Genomes European phase I (a) panel (**Online Methods**). Finally, we performed a joint analysis of the discovery and replication phases.

We identified 97 statistically independent SNPs meeting replication criteria ($p_{replication} < 0.05$, $p_{joint} < 5 \times 10^{-8}$, and $p_{joint} < p_{discovery}$); 93 primary signals (Supplementary Figs. 3-95) and four secondary signals. Of these, 48 are novel to multiple sclerosis (**Table 1**) and 49 correspond to previously identified multiple sclerosis effects (**Table 2**). An additional 11 independent SNPs showed suggestive evidence of association ($p_{joint} < 1 \times 10^{-6}$) (**Supplementary Table 2**).

The strongest novel association, rs12087340 ($p_{joint} = 1.1 \times 10^{-20}$, OR = 1.21), lies between *BCL10* (B-cell CLL / lymphoma 10) and *DDAH1* (dimethylarginine dimethylaminohylaminohydrolase 1). The protein encoded by *BCL10* contains a caspase recruitment domain (CARD) and has been shown to activate NF-kappaB.[17] The latter is a signalling molecule that plays an important role in controlling gene expression in inflammation, immunity, cell proliferation, and apoptosis. It has been pursued as a potential therapeutic target for multiple sclerosis.[18] *BCL10* is also reported to interact with other CARD domain containing proteins including *CARD11*.[19] We have also identified a novel association of rs1843938 ($p_{joint} = 1.2 \times 10^{-10}$, OR = 1.08), which is only 30 kb from *CARD11*.

One novel SNP was found within an exon, rs2288904 ($p_{joint} = 1.6 \times 10^{-11}$, OR= 1.10); a missense variant in *SLC44A2* (solute carrier family 44, member 2). Notably, this variant is also reported as a monocyte-specifc cis-acting eQTL for the antisense transcript of the nearby *ILF3* (interleukin enhancer binding factor 3).[20] This protein was first discovered to be a subunit of a nuclear factor found in activated T-cells, which is required for T-cell expression of *IL2*, an important molecule regulating many aspects of inflammation.

Of the 49 previously identified effects,[9,10,21] 37 are in designated fine-mapping regions, and 23 of these 37 signals were localized to a single gene based on genomic position (Supplementary Table 3). Recognizing that proximity does not necessarily indicate functional importance, this emphasizes the utility of dense mapping in localizing signals from a genome-wide screen. The ImmunoChip analysis furthered the understanding of previously proposed secondary signals at three loci (**Supplementary Note** and **Supplementary Tables 4-6**); in particular we showed that the effects of two previously proposed independent associations at the *IL2RA* locus are driven by a single variant, rs2104286.[7,22].

In an effort to define the functionally relevant variants underlying these associations, we further studied the regions surrounding the 97 associated SNPs using both a Bayesian and frequentist approach in 6,356 multiple sclerosis subjects and 9,617 healthy controls from the UK (**Online Methods**).[23] Based on imputation quality, fine-mapping was possible in 68 regions (**Supplementary Table 7**): 66 of 93 primary (**Fig. 2A**) and two of four secondary. Eight of the 68 regions were fine-mapped to high resolution (**Table 3**, **Fig. 2B** and **Supplementary Fig. 96**). One third of the variants identified in these eight regions were imputed, indicating reliance on imputation even with dense genotyping coverage.

To assess whether functional annotation[24] provides clues about the molecular mechanisms associated with genetic risk, we considered the relationship of variants to described coding and regulatory features in these eight regions. Although we found no variants with missense or nonsense effects, there was a notable enrichment for variants with functional effects: one known to affect splicing,[25] three known to correlate with RNA or serum protein levels[22,26,27] and several in transcription-factor binding and DNase I hypersensitive sites.[28,29] Four of the 18 variants in the fine-mapped regions are within conserved regions (GERP > 2).[30] This lack of functional annotation likely reflects the limited repertoire of reference expression and epigenomic profiles and suggests that the function of the variants may be cell-type or cell-state specific, as has been reported for many eQTLs in immune cell types.[20]

To determine the Gene Ontology (GO) processes of the 97 associated variants, we used MetaCore from Thomson Reuters (**Online Methods**). We found the majority of the 97 variants lie within 50 kb of genes having immunological function. Of the 86 unique genes represented, 35 are linked to the GO immune system process (**Table 1** and **Table 2**). We do not see a substantial over- or under- representation of certain GO processes when comparing the novel and previously identified loci, but this may be a limitation of ImmunoChip targeting genomic loci enriched for immunologically active genes, with more subtle distinctions between them not adequately captured by broad annotations such as GO.

Finally, we explored the overlap between our findings and those in autoimmune diseases with reported ImmunoChip analyses. We calculated the percentage of multiple sclerosis signals (110 non-MHC, **Supplementary Table 8**) overlapping those of other autoimmune diseases by requiring an $r^2 \geq 0.8$ between the best variants reported in each study using SNAP.[31] In total we find that ~22% of our signals overlap at least one other autoimmune disease. More specificially, ~9.1% overlap with inflammatory bowel disease (IBD) - ~7.3% with ulcerative colitis (UC), ~9.1% with Crohn's disease (CD) - [15], ~9.1% with primary biliary cirrhosis (PBC),[32,33] ~4.5% with celiac disease (CeD),[34] ~4.5% with rheumatoid arthritis (RhA),[35] ~0.9% with psoriasis (PS),[36] and ~2.7% with autoimmune thyroid disease (AITD).[37] We report the same top variant seen in PBC for 7 loci. We also note that our best *TYK2* variant (rs34536443)[38] is also the most

associated variant for PBC, PS and RhA. Lastly, AITD, CeD, PBC, and RhA report variants with pairwise $r^2 \geq 0.8$ to the multiple sclerosis variant near *MMEL1*[39] (**Supplementary Table 8**).

In summary, we have identified 48 new multiple sclerosis susceptibility variants. These novel loci expand our understanding of the immune system processes implicated in multiple sclerosis. We estimate that the 110 non-MHC established risk variants explain 20% of the sibling recurrence risk; 28% including the already identified MHC effects[9] (**Supplementary Note**). Additionally, we have identified five regions where consistent high resolution fine-mapping implicated one variant which accounted for more than 50% of the posterior in previously identified regions of *TNFSF14*, *IL2RA*, *TNFRSF1A*, *IL12A*, and *STAT4*. Our study further implicates NF-kappaB in multiple sclerosis pathobiology[18], emphasizes the value of dense fine-mapping in large follow-up data sets, and exposes the urgent need for functional annotation in relevant tissues. Understanding the implicated networks and their relation to environmental risk factors will promote the development of rational therapies and may enable the development of preventive strategies.

URLs

ImmunoBase, http://www.immunobase.org/; eQTL browser, http://eqtl.uchicago.edu/; MetaCore, https://portal.genego.com/.

Figure Legends

Figure 1. Discovery phase results. Primary association analysis of 161,311 autosomal variants in the discovery phase (based on 14,498 cases and 24,091 healthy controls). The outer most track shows the numbered autosomal chromosomes. The second track indicates the gene closest to the most associated SNP meeting all replication criteria. Previously identified associations are indicated in grey. The third track indicates the physical position of the 184 fine-mapping intervals (green). The inner most track indicates –log(p) (two-sided) for each SNP (scaled from 0-12 which truncates the signal in several regions, see Supplementary Table 1). Additionally, contour lines are given at the *a priori* discovery(–log(p) = 4) and genome-wide significance (-log(p) = 7.3) thresholds. Orange indicates -log(p) $\geq$ 4 and < 7.3, while red indicates –log(p) $\geq$ 7.3. Details of the full discovery phase results can be found in ImmunoBase.

Figure 2. Bayesian fine-mapping within primary regions of association. a) Summary of the extent of fine-mapping across 66 regions in 9,617 healthy controls from the UK, showing the the physical extent of, the number of variants, and the number of genes spanned by the posterior 90% and 50% credible sets. b) Detail of fine-mapping in region of *TNFSF14*. Above the x-axis indicates the Bayes Factor summarizing evidence for association for the SNPs prior to conditioning (blue markers) while below the x-axis indicates the Bayes Factor after conditioning on the lead SNP (rs1077667). Mb=Megabases.

Table 1 48 Novel non-MHC susceptibility loci associated with multiple sclerosis at a genome-wide significance level

				Discovery			Replication			Joint			
SNP	**Chr**	**Position[a]**	**RA**	**RAF**	**P-value**	**OR**	**RAF**	**P-value**	**OR**	**P-value**	**OR**	**Gene[b]**	**Function**
rs3007421	1	6530189	A	0.12	9.6×10^{-7}	1.12	0.13	8.8×10^{-5}	1.10	4.7×10^{-10}	1.11	*PLEKHG5*	intronic
rs12087340	1	85746993	A	0.09	5.1×10^{-12}	1.22	0.09	2.9×10^{-10}	1.20	1.1×10^{-20}	1.21	***BCL10***	intergenic
rs11587876	1	85915183	A	0.79	8.4×10^{-8}	1.12	0.81	2.9×10^{-3}	1.06	4.4×10^{-9}	1.09	*DDAH1*	intronic
rs666930	1	120258970	G	0.53	7.5×10^{-8}	1.09	0.53	1.3×10^{-5}	1.07	6.0×10^{-12}	1.08	*PHGDH*	intronic
rs2050568	1	157770241	G	0.53	1.3×10^{-6}	1.08	0.54	2.3×10^{-5}	1.07	1.5×10^{-10}	1.08	*FCRL1*	intronic
rs35967351	1	160711804	A	0.67	1.7×10^{-6}	1.09	0.68	5.9×10^{-6}	1.09	4.4×10^{-11}	1.09	***SLAMF7***	intronic
rs4665719	2	25017860	G	0.25	6.8×10^{-6}	1.09	0.25	1.1×10^{-4}	1.08	3.1×10^{-9}	1.08	*CENPO*	intronic
rs842639	2	61095245	A	0.65	1.7×10^{-9}	1.11	0.67	1.4×10^{-6}	1.09	2.0×10^{-14}	1.10	*FLJ16341*	ncRNA
rs9967792	2	191974435	G	0.62	1.8×10^{-9}	1.11	0.64	1.2×10^{-4}	1.07	3.5×10^{-12}	1.09	***STAT4***	intronic
rs11719975	3	18785585	C	0.27	5.4×10^{-6}	1.09	0.28	4.1×10^{-4}	1.07	1.1×10^{-8}	1.08		intergenic
rs4679081	3	33013483	G	0.52	1.2×10^{-5}	1.08	0.55	3.7×10^{-4}	1.07	2.2×10^{-9}	1.07	***CCR4***	intergenic
rs9828629	3	71530346	G	0.62	5.5×10^{-6}	1.08	0.64	8.5×10^{-6}	1.08	1.9×10^{-10}	1.08	***FOXP1***	intronic
rs2726518	4	106173199	C	0.55	1.2×10^{-5}	1.09	0.58	4.7×10^{-4}	1.06	3.9×10^{-8}	1.07	***TET2***	intronic
rs756699	5	133446575	A	0.87	3.0×10^{-6}	1.12	0.88	6.5×10^{-6}	1.11	8.8×10^{-11}	1.12	***TCF7***	intergenic
none[c]	5	141506564	C	0.61	6.0×10^{-5}	1.07	0.62	1.5×10^{-5}	1.08	3.6×10^{-9}	1.07	*NDFIP1*	intronic
rs4976646	5	176788570	G	0.34	1.0×10^{-12}	1.13	0.36	5.0×10^{-7}	1.10	4.4×10^{-18}	1.12	*RGS14*	intronic
rs17119	6	14719496	A	0.81	1.9×10^{-6}	1.11	0.80	1.2×10^{-5}	1.10	1.0×10^{-10}	1.10		intergenic
rs941816	6	36375304	G	0.18	4.5×10^{-9}	1.13	0.20	8.3×10^{-5}	1.08	3.9×10^{-12}	1.11	*PXT1*	intronic
rs1843938	7	3113034	A	0.44	2.2×10^{-6}	1.08	0.44	1.1×10^{-5}	1.08	1.2×10^{-10}	1.08	***CARD11***	intergenic
rs706015	7	27014988	C	0.18	1.3×10^{-9}	1.14	0.18	9.9×10^{-3}	1.06	1.1×10^{-9}	1.10		intergenic
rs917116	7	28172739	C	0.20	2.1×10^{-8}	1.12	0.21	5.8×10^{-3}	1.06	3.3×10^{-9}	1.09	*JAZF1*	intronic
rs60600003	7	37382465	C	0.10	2.5×10^{-8}	1.16	0.10	4.2×10^{-7}	1.14	6.0×10^{-14}	1.15	*ELMO1*	intronic
rs201847125[d]	7	50325567	G	0.70	2.9×10^{-8}	1.11	0.70	6.7×10^{-5}	1.09	1.2×10^{-11}	1.10	***IKZF1***	intergenic
rs2456449	8	128192981	G	0.36	2.2×10^{-8}	1.10	0.37	3.8×10^{-3}	1.05	1.8×10^{-9}	1.08		intergenic
rs793108	10	31415106	A	0.50	5.6×10^{-8}	1.09	0.51	1.8×10^{-5}	1.07	6.1×10^{-12}	1.08		intergenic
rs2688608	10	75658349	A	0.55	6.4×10^{-5}	1.07	0.56	2.0×10^{-4}	1.06	4.6×10^{-8}	1.07	*C10orf55*	intergenic
rs7120737	11	47702395	G	0.15	7.6×10^{-8}	1.13	0.15	1.0×10^{-3}	1.08	1.0×10^{-9}	1.10	*AGBL2*	intronic
rs694739	11	64097233	A	0.62	1.3×10^{-5}	1.08	0.62	3.8×10^{-5}	1.07	2.0×10^{-9}	1.07	***PRDX5***	intergenic
rs9736016	11	118724894	T	0.63	2.2×10^{-8}	1.10	0.63	2.6×10^{-8}	1.10	3.0×10^{-15}	1.10	***CXCR5***	intergenic
rs12296430	12	6503500	C	0.19	3.6×10^{-10}	1.14	0.21	1.7×10^{-5}	1.09	7.2×10^{-14}	1.12	***LTBR***	intergenic
rs4772201	13	100086259	A	0.82	1.7×10^{-7}	1.12	0.83	1.1×10^{-4}	1.09	1.3×10^{-10}	1.10	*MIR548AN*	intergenic
rs12148050	14	103263788	A	0.35	1.5×10^{-5}	1.08	0.36	4.3×10^{-9}	1.10	5.1×10^{-13}	1.09	***TRAF3***	intronic
rs59772922	15	79207466	A	0.83	4.0×10^{-6}	1.11	0.83	5.4×10^{-4}	1.08	1.2×10^{-8}	1.09	***CTSH***	intergenic
rs8042861	15	90977333	A	0.44	9.8×10^{-7}	1.08	0.45	3.4×10^{-4}	1.06	2.2×10^{-9}	1.07	*IQGAP1*	intronic
rs6498184	16	11435990	G	0.81	2.1×10^{-10}	1.15	0.82	6.5×10^{-9}	1.14	7.4×10^{-18}	1.15	*RMI2*	intergenic
rs7204270[*]	16	30156963	G	0.50	9.3×10^{-8}	1.09	0.49	3.7×10^{-5}	1.08	1.6×10^{-11}	1.09	***MAPK3***	intergenic
rs1886700	16	68685905	A	0.14	8.8×10^{-6}	1.11	0.14	3.2×10^{-4}	1.08	1.3×10^{-8}	1.10	*CDH3*	intronic
rs12149527	16	79110596	A	0.47	1.7×10^{-6}	1.08	0.47	4.3×10^{-6}	1.08	3.3×10^{-11}	1.08	*WWOX*	intronic
rs7196953	16	79649394	A	0.29	2.6×10^{-5}	1.08	0.30	7.9×10^{-7}	1.09	1.0×10^{-10}	1.09	*MAF*	intergenic
rs12946510	17	37912377	A	0.47	8.5×10^{-6}	1.08	0.48	8.0×10^{-5}	1.07	2.9×10^{-9}	1.07	***IKZF3***	intergenic
rs4794058	17	45597098	A	0.50	1.6×10^{-5}	1.07	0.52	3.5×10^{-10}	1.11	1.0×10^{-13}	1.09	***NPEPPS***	intergenic
rs2288904	19	10742170	G	0.77	9.6×10^{-10}	1.14	0.78	5.4×10^{-4}	1.07	1.6×10^{-11}	1.10	*SLC44A2*	exonic
rs1870071	19	16505106	G	0.29	5.7×10^{-10}	1.12	0.30	4.6×10^{-7}	1.09	2.0×10^{-15}	1.10	*EPS15L1*	intronic
rs17785991	20	48438761	A	0.35	6.4×10^{-7}	1.09	0.34	5.9×10^{-3}	1.05	4.2×10^{-8}	1.07	*SLC9A8*	intronic
rs2256814	20	62373983	A	0.19	8.3×10^{-7}	1.11	0.21	6.4×10^{-4}	1.08	3.5×10^{-9}	1.09	*SLC2A4RG*	intronic
Secondary													
rs7769192[e]	6	137962655	G	0.55	1.3×10^{-5}	1.08	0.54	5.1×10^{-5}	1.07	3.3×10^{-9}	1.08		intergenic
rs533646[f]	11	118566746	G	0.68	3.6×10^{-7}	1.10	0.68	3.9×10^{-5}	1.08	7.6×10^{-11}	1.09	*TREH*	intergenic
rs4780346[g]	16	11288806	A	0.23	6.8×10^{-6}	1.09	0.25	1.5×10^{-5}	1.09	4.4×10^{-10}	1.09	*CLEC16A*	intergenic

All listed signals had a discovery P-value $\leq 1.0 \times 10^{-4}$, a replication P-value $\leq 5.0 \times 10^{-2}$, and a joint P-value $\leq 5.0 \times 10^{-8}$

All P-values are two-sided

RA= Risk Allele, RAF = Risk Allele Frequency

[a]Position is based on human genome 19 and dbSNP 137.

[b]Nearest gene listed if within 50Kb. Bold indicates Gene Ontology Immune System Process.

A proxy SNP [c](rs1036207, $r^2 = 0.99$) and [d](rs716719, $r^2=1.00$) was used in replication.

The P-value and OR values provided are after conditioning on [e]rs67297943 (Known – see Table 2), [f]rs9736016, and [g]rs12927355 (Known – see Table 2).

[*]Note primary was rs11865086 (P-value = 1.77×10^{-8}) in Discovery but not available in Replication so the best proxy was used.

Table 2 49 Known non-MHC susceptibility loci associated with multiple sclerosis at a genome-wide significance level

				Discovery			Replication			Joint			
SNP	**Chr**	**Position[a]**	**RA**	**RAF**	**P-value**	**OR**	**RAF**	**P-value**	**OR**	**P-value**	**OR**	**Gene[b]**	**Function**
rs3748817	1	2525665	A	0.64	1.3×10^{-12}	1.14	0.65	1.2×10^{-15}	1.15	1.3×10^{-26}	1.14	*MMEL1*	intronic
rs41286801	1	92975464	A	0.14	7.9×10^{-16}	1.20	0.16	2.1×10^{-12}	1.17	1.4×10^{-26}	1.19	*EVI5*	UTR3
rs7552544*	1	101240893	A	0.56	3.7×10^{-6}	1.08	0.43	3.3×10^{-12}	1.12	1.9×10^{-16}	1.10	***VCAM1***	intergenic
rs6677309	1	117080166	A	0.88	1.5×10^{-28}	1.34	0.88	4.1×10^{-16}	1.24	5.4×10^{-42}	1.29	***CD58***	intronic
rs1359062	1	192541472	C	0.82	1.8×10^{-13}	1.18	0.83	2.1×10^{-8}	1.13	4.8×10^{-20}	1.15	***RGS1***	intergenic
rs55838263	1	200874728	A	0.71	1.4×10^{-9}	1.12	0.71	3.9×10^{-11}	1.13	4.0×10^{-19}	1.13	*C1orf106*	intronic
rs2163226	2	43361256	A	0.71	7.0×10^{-8}	1.10	0.73	3.8×10^{-10}	1.14	2.1×10^{-16}	1.12		intergenic
rs7595717	2	68587477	A	0.26	3.3×10^{-7}	1.10	0.27	6.8×10^{-8}	1.10	1.2×10^{-13}	1.10	***PLEK***	intergenic
rs9989735	2	231115454	C	0.18	7.8×10^{-14}	1.17	0.19	6.8×10^{-11}	1.14	4.2×10^{-23}	1.16	*SP140*	intronic
rs2371108	3	27757018	A	0.38	2.1×10^{-6}	1.08	0.39	5.8×10^{-11}	1.12	1.5×10^{-15}	1.10	***EOMES***	downstream
rs1813375	3	28078571	A	0.47	5.7×10^{-18}	1.15	0.49	4.4×10^{-16}	1.15	1.9×10^{-32}	1.15		intergenic
rs1131265	3	119222456	C	0.80	2.0×10^{-15}	1.19	0.81	4.8×10^{-10}	1.14	1.4×10^{-23}	1.17	*TIMMDC1*	exonic
rs1920296*	3	121543577	C	0.64	6.8×10^{-15}	1.14	0.64	5.5×10^{-9}	1.10	6.5×10^{-22}	1.12	*IQCB1*	intronic
rs2255214*	3	121770539	C	0.52	5.3×10^{-13}	1.13	0.52	3.3×10^{-13}	1.13	1.2×10^{-24}	1.13	***CD86***	intergenic
rs1014486	3	159691112	G	0.43	1.2×10^{-9}	1.11	0.44	1.4×10^{-10}	1.11	1.1×10^{-18}	1.11	***IL12A***	intergenic
rs7665090	4	103551603	G	0.52	2.4×10^{-6}	1.08	0.53	5.0×10^{-4}	1.13	1.0×10^{-8}	1.09	*MANBA*	intergenic
rs6881706	5	35879156	C	0.72	4.9×10^{-9}	1.12	0.73	1.7×10^{-9}	1.12	4.3×10^{-17}	1.12	***IL7R***	intergenic
rs6880778	5	40399096	G	0.60	1.7×10^{-8}	1.10	0.61	3.9×10^{-13}	1.13	8.1×10^{-20}	1.12		intergenic
rs71624119	5	55440730	G	0.76	2.7×10^{-9}	1.12	0.76	1.9×10^{-5}	1.09	3.4×10^{-13}	1.11	*ANKRD55*	intronic
rs72928038	6	90976768	A	0.17	7.6×10^{-7}	1.11	0.19	9.0×10^{-11}	1.17	1.5×10^{-15}	1.14	*BACH2*	intronic
rs11154801	6	135739355	A	0.37	2.3×10^{-9}	1.11	0.37	1.0×10^{-12}	1.13	1.8×10^{-20}	1.12	*AHI1*	intronic
rs17066096	6	137452908	G	0.23	5.9×10^{-12}	1.14	0.25	4.1×10^{-13}	1.15	1.6×10^{-23}	1.14	*IL22RA2*	intergenic
rs67297943	6	138244816	A	0.78	4.8×10^{-8}	1.12	0.80	2.5×10^{-6}	1.11	5.5×10^{-13}	1.11	***TNFAIP3***	intergenic
rs212405	6	159470559	T	0.62	1.4×10^{-15}	1.15	0.64	1.8×10^{-7}	1.10	8.0×10^{-21}	1.12	*TAGAP*	intergenic
rs1021156	8	79575804	A	0.24	5.6×10^{-10}	1.12	0.26	2.1×10^{-8}	1.11	8.5×10^{-17}	1.11	*ZC2HC1A*	intergenic
rs4410871	8	128815029	G	0.72	2.0×10^{-9}	1.12	0.72	3.4×10^{-8}	1.11	4.3×10^{-16}	1.11	*MIR1204*	intergenic
rs759648	8	129158945	C	0.31	2.8×10^{-6}	1.09	0.31	3.7×10^{-5}	1.08	5.0×10^{-10}	1.08	*MIR1208*	intergenic
rs2104286	10	6099045	A	0.72	7.6×10^{-23}	1.21	0.73	3.6×10^{-26}	1.23	2.3×10^{-47}	1.22	***IL2RA***	intronic
rs1782645	10	81048611	A	0.43	4.3×10^{-7}	1.09	0.41	6.2×10^{-10}	1.11	2.5×10^{-15}	1.10	*ZMIZ1*	intronic
rs7923837	10	94481917	G	0.61	4.6×10^{-9}	1.11	0.62	2.0×10^{-9}	1.11	4.3×10^{-17}	1.11	***HHEX***	intergenic
rs34383631	11	60793330	A	0.40	5.7×10^{-10}	1.11	0.39	4.5×10^{-15}	1.15	3.7×10^{-23}	1.13	*CD6*	intergenic
rs1800693	12	6440009	G	0.40	6.9×10^{-16}	1.14	0.41	1.0×10^{-13}	1.14	6.7×10^{-28}	1.14	*TNFRSF1A*	intronic
rs11052877	12	9905690	G	0.36	5.4×10^{-9}	1.10	0.38	1.2×10^{-5}	1.08	5.6×10^{-13}	1.09	***CD69***	UTR3
rs201202118[c]	12	58182062	A	0.67	7.4×10^{-13}	1.14	0.67	1.6×10^{-10}	1.12	9.0×10^{-22}	1.13	*TSFM*	intronic
rs7132277	12	123593382	A	0.19	1.9×10^{-6}	1.10	0.19	1.4×10^{-8}	1.13	1.9×10^{-13}	1.12	*PITPNM2*	intronic
rs2236262	14	69261472	A	0.50	1.2×10^{-5}	1.08	0.50	3.8×10^{-8}	1.09	2.5×10^{-12}	1.08	***ZFP36L1***	intronic
rs74796499	14	88432328	C	0.95	8.5×10^{-11}	1.31	0.95	4.5×10^{-11}	1.33	2.4×10^{-20}	1.32	*GALC*	intronic
rs12927355	16	11194771	G	0.68	8.2×10^{-27}	1.21	0.69	4.3×10^{-21}	1.18	6.4×10^{-46}	1.20	*CLEC16A*	intronic
rs35929052	16	85994484	G	0.89	3.3×10^{-7}	1.14	0.88	3.6×10^{-6}	1.15	5.9×10^{-12}	1.15	***IRF8***	intergenic
rs4796791	17	40530763	A	0.36	1.8×10^{-8}	1.10	0.36	1.2×10^{-13}	1.14	3.7×10^{-20}	1.12	*STAT3*	intronic
rs8070345	17	57816757	A	0.45	5.4×10^{-16}	1.14	0.46	1.9×10^{-9}	1.10	2.2×10^{-23}	1.12	*VMP1*	intronic
rs1077667	19	6668972	G	0.79	3.5×10^{-13}	1.16	0.79	8.4×10^{-13}	1.16	1.7×10^{-24}	1.16	***TNFSF14***	intronic
rs34536443	19	10463118	C	0.95	1.2×10^{-8}	1.28	0.96	2.9×10^{-7}	1.30	1.8×10^{-14}	1.29	***TYK2***	exonic
rs11554159	19	18285944	G	0.73	2.6×10^{-13}	1.15	0.74	1.4×10^{-12}	1.15	1.9×10^{-24}	1.15	***IFI30***	exonic
rs8107548	19	49870643	G	0.25	2.0×10^{-6}	1.09	0.26	2.5×10^{-10}	1.13	5.7×10^{-15}	1.11	*DKKL1*	intronic
rs4810485	20	44747947	A	0.25	1.8×10^{-5}	1.08	0.25	1.4×10^{-12}	1.14	7.7×10^{-16}	1.11	***CD40***	intronic
rs2248359	20	52791518	G	0.60	9.8×10^{-5}	1.07	0.62	8.2×10^{-11}	1.12	2.0×10^{-13}	1.09	*CYP24A1*	intergenic
rs2283792	22	22131125	C	0.51	1.1×10^{-6}	1.08	0.53	5.4×10^{-11}	1.11	5.5×10^{-16}	1.10	***MAPK1***	intronic
Secondary													
rs523604[d]	11	118755738	A	0.53	2.5×10^{-7}	1.09	0.54	4.0×10^{-9}	1.11	6.2×10^{-15}	1.10	***CXCR5***	intronic

All listed signals had a discovery P-value $\leq 1.0 \times 10^{-4}$, a replication P-value $\leq 5.0 \times 10^{-2}$, and a joint P-value $\leq 5.0 \times 10^{-8}$

All P-values are two-sided

RA = Risk Allele, RAF = Risk Allele Frequency

[a]Position is based on human genome 19 and dbSNP 137.

[b]Nearest gene listed if within 50Kb. Bold indicates Gene Ontology Immune System Process.

[c]A proxy SNP (rs10431552, $r^2 = 0.99$) was used in replication.

[d]The P-value and OR values provided are after conditioning on rs9736016 and rs533646 (both Novel – see Table 1).

* These three SNPs were not primary in the 2011 GWAS, two were secondary and the third was tertiary in that study.

Table 3 The 18 variants from the 8 regions with consistent high resolution fine-mapping

Gene	SNP	Chr	Position[a]	Posterior	GERP	Functional Annotation[b]
TNFSF14	rs1077667	19	6668972	0.74	-3.89	intronic, TFBS / DNase1 peak, correlates with serum levels of TNFSF14
IL2RA	rs2104286	10	6099045	0.93	-0.47	intronic, correlates with soluble IL-2RA levels
TNFRSF1A	rs1800693	12	6440009	0.69	2.53	intronic, causes splicing defect and truncated soluble TNFRSF1A
	rs4149580[c]	12	6446990	0.10	2.06	intronic
IL12A	rs1014486	3	159691112	0.67	0.24	-
CD6	rs34383631	11	60793330	0.20	1.66	-
	rs4939490[c]	11	60793651	0.14	-0.53	-
	rs4939491[c]	11	60793722	0.14	-0.37	-
	rs4939489	11	60793648	0.10	3.25	-
TNFAIP3	rs632574	6	137959118	0.27	-1.15	-
	rs498549[c]	6	137984935	0.20	0.52	-
	rs651973	6	137996134	0.17	2.41	downstream of RP11-95M15.1 lincRNA gene
	rs536331	6	137993049	0.15	0.19	upstream of RP11-95M15.1 lincRNA gene
CD58	rs6677309	1	117080166	0.21	-1.18	intronic, TFBS / DNase1 peak
	rs35275493[c]	1	117095502	0.24	0.75	intronic (insertion)
	rs10754324[c]	1	117093035	0.22	0.32	intronic
	rs1335532	1	117100957	0.17	-1.32	intronic
STAT4	rs78712823	2	191958581	0.59	-3.98	intronic

All listed variants have posterior ≥ 0.1 in regions where ≤ 5 variants explain the top 50% of the posterior and the top SNP from the frequentist analysis lives in the 90% confidence interval, ordered by maximum posterior.

Posterior denotes the posterior probability of any variant driving association. GERP denotes Genomic Evolutionary Rate Profiling.

[a]Position is based on human genome 19 and dbSNP 137.

[b]Functional data from VEP, eQTL browser, Fairfax et al. (2012), pubmed searches, 1000G. Dash indicates intergenic with no additional annotation. Variants without annotation are intergenic and have no reported regulatory consequence.

[c]Imputed variant.

Acknowledgements

We thank participants, referring nurses, physicians, and funders - National Institutes of Health, Wellcome Trust, UK MS Society, UK Medical Research Council, US National MS Society, Cambridge NIHR BRC, DeNDRon, Bibbi and Niels Jensens Foundation, Swedish Brain Foundation, Swedish Research Council, Knut and Alice Wallenberg Foundation, Swedish Heart-Lung Foundation, Foundation for Strategic Research, Stockholm County Council, Karolinska Institutet, Institut National de la Santé et de la Recherche Médicale, Fondation d'Aide pour la Recherche sur la Sclérose En Plaques, Association Française contre les Myopathies, GIS-IBISA, German Ministry for Education and Research, German Competence Network MS, Deutsche Forschungsgemeinschaft, Munich Biotec Cluster M4, Fidelity Biosciences Research Initiative, Research Foundation Flanders, Research Fund KU Leuven, Belgian Charcot Foundation, Gemeinnützige Hertie Stiftung, University Zurich, Danish MS Society, Danish Council for Strategic Research, Academy of Finland, Sigrid Juselius Foundation, Helsinki University, Italian MS Foundation, Fondazione Cariplo, Italian Ministry of University and Research, CRT Foundation of Turin, Italian Ministry of Health, Italian Institute of Experimental Neurology, MS association of Oslo, Norwegian Research Council, South Eastern Norwegian Health Authorities, Australian National Health and Medical Research Council, Dutch MS Foundation, Kaiser Permanente. We acknowledge British 1958 Birth Cohort, UK National Blood Service, Vanderbilt University Medical Center's BioVU DNA Resources Core, Centre de Ressources Biologiques du Réseau Français d'Etude Génétique de la Sclérose en Plaques, Norwegian Bone Marrow Registry, Norwegian MS Registry and Biobank, North American Research Committee on MS Registry, Brigham and Womens Hospital PhenoGenetic Project and DILGOM funded by the Academy of Finland. See Supplementary Note for details.

Author List and Affiliations

Ashley H Beecham[1, 126], Nikolaos A Patsopoulos[2-6, 126], Dionysia K Xifara[7], Mary F Davis[8], Anu Kemppinen[9], Chris Cotsapas[10-12], Tejas S Shahi[13], Chris Spencer[7], David Booth[14], An Goris[15], Annette Oturai[16], Janna Saarela[17], Bertrand Fontaine[18], Bernhard Hemmer[19-21], Claes Martin[22], Frauke Zipp[23], Sandra D'alfonso[24], Filippo Martinelli-Boneschi[25, 26], Bruce Taylor[27], Hanne F Harbo[28, 29], Ingrid Kockum[30], Jan Hillert[30], Tomas Olsson[30], Maria Ban[9], Jorge R Oksenberg[31], Rogier Hintzen[32], Lisa F Barcellos[33-35], Wellcome Trust Case Control Consortium 2 (WTCCC2)[36], International IBD Genetics Consortium (IIBDGC)[36], Cristina Agliardi[37], Lars Alfredsson[38], Mehdi Alizadeh[39], Carl Anderson[13], Robert Andrews[13], Helle Bach Søndergaard[16], Amie Baker[9], Gavin Band[7], Sergio E Baranzini[31], Nadia Barizzone[24], Jeffrey Barrett[13], Céline Bellenguez[7], Laura Bergamaschi[24], Luisa Bernardinelli[40], Achim Berthele[19], Viola Biberacher[19], Thomas M C Binder[41], Hannah Blackburn[13], Izaura L Bomfim[30], Paola Brambilla[25], Simon Broadley[42], Bruno Brochet[43], Lou Brundin[30], Dorothea Buck[19], Helmut Butzkueven[44, 45], Stacy J Caillier[31], William Camu[46], Wassila Carpentier[47], Paola Cavalla[48, 49], Elisabeth G Celius[28], Irène Coman[50], Giancarlo Comi[25, 26], Lucia Corrado[24], Leentje Cosemans[15], Isabelle Cournu-Rebeix[18], Bruce A C Cree[31], Daniele Cusi[51], Vincent Damotte[18], Gilles Defer[52], Silvia R Delgado[53], Panos Deloukas[13], Alessia di Sapio[54], Alexander T Dilthey[7], Peter Donnelly[7], Bénédicte Dubois[15], Martin Duddy[55], Sarah Edkins[13], Irina Elovaara[56], Federica Esposito[25, 26], Nikos Evangelou[57], Barnaby Fiddes[9], Judith Field[58], Andre Franke[59], Colin Freeman[7], Irene Y Frohlich[2], Daniela Galimberti[60, 61], Christian Gieger[62], Pierre-Antoine Gourraud[31], Christiane Graetz[23], Andrew Graham[63], Verena Grummel[19], Clara Guaschino[25, 26], Athena Hadjixenofontos[1], Hakon Hakonarson[64, 65], Christopher Halfpenny[66], Gillian Hall[67], Per Hall[68], Anders Hamsten[69], James Harley[70], Timothy Harrower[71], Clive Hawkins[72], Garrett Hellenthal[73], Charles Hillier[74], Jeremy Hobart[75], Muni Hoshi[19], Sarah E Hunt[13], Maja Jagodic[30], Ilijas Jelčić[76, 77], Angela Jochim[19], Brian Kendall[78], Allan Kermode[79, 80], Trevor Kilpatrick[81], Keijo Koivisto[82], Ioanna Konidari[1], Thomas Korn[19], Helena Kronsbein[19], Cordelia Langford[13], Malin Larsson[83], Mark Lathrop[84-86], Christine Lebrun-Frenay[87], Jeannette Lechner-Scott[88], Michelle H Lee[2], Maurizio A Leone[89], Virpi Leppä[17], Giuseppe Liberatore[25, 26], Benedicte A Lie[29, 90], Christina M Lill[23, 91], Magdalena Lindén[30], Jenny Link[30], Felix Luessi[23], Jan Lycke[92], Fabio Macciardi[93, 94], Satu Männistö[95], Clara P Manrique[1], Roland Martin[76, 77], Vittorio Martinelli[26], Deborah Mason[96], Gordon Mazibrada[97], Cristin McCabe[10], Inger-Lise Mero[28, 29, 90], Julia Mescheriakova[32], Loukas Moutsianas[7], Kjell-Morten Myhr[98], Guy Nagels[99], Richard Nicholas[100], Petra Nilsson[101], Fredrik Piehl[30], Matti Pirinen[7], Siân E Price[102], Hong Quach[33], Mauri Reunanen[103, 104], Wim Robberecht[105-107], Neil P Robertson[108], Mariaemma Rodegher[26], David Rog[109], Marco Salvetti[110], Nathalie C Schnetz-Boutaud[8], Finn Sellebjerg[16], Rebecca C Selter[19], Catherine Schaefer[35], Sandip Shaunak[111], Ling Shen[35], Simon Shields[112], Volker Siffrin[23], Mark Slee[113], Per Soelberg Sorensen[16], Melissa Sorosina[25], Mireia Sospedra[76, 77], Anne Spurkland[114], Amy Strange[7], Emilie Sundqvist[30], Vincent Thijs[105-107], John Thorpe[115], Anna Ticca[116], Pentti Tienari[117], Cornelia van Duijn[118], Elizabeth M Visser[119], Steve Vucic[14], Helga Westerlind[30], James S Wiley[58], Alastair Wilkins[120], James F Wilson[121], Juliane Winkelmann[19, 20, 122, 123], John Zajicek[75], Eva Zindler[23], Jonathan L Haines[8], Margaret A Pericak-Vance[1], Adrian J Ivinson[124], Graeme Stewart[14], David Hafler[10, 11, 125], Stephen L Hauser[31], Alastair Compston[9], Gil McVean[7], Philip De Jager[2, 5, 10, 126], Stephen Sawcer[9, 126], Jacob L McCauley[1, 126]

1. John P. Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, Miami, Florida, USA.
2. Department of Neurology, Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Brigham & Women's Hospital, Boston, Massachusetts, USA.
3. Department of Psychiatry, Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Brigham & Women's Hospital, Boston, Massachusetts, USA.
4. Department of Medicine, Division of Genetics, Brigham & Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA.
5. Harvard Medical School, Boston, Massachusetts, USA.
6. Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
7. The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
8. Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, Tennessee, USA.
9. Department of Clinical Neurosciences, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK.
10. Program in Medical and Population Genetics, Broad Institute of Harvard University and MIT, Cambridge, Massachusetts, USA.
11. Department of Neurology, Yale University School of Medicine, New Haven, Connecticut, USA.
12. Department of Genetics, Yale University School of Medicine, New Haven, Connecticut, USA.
13. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK.
14. Westmead Millennium Institute, University of Sydney, New South Wales, Australia.
15. Section of Experimental Neurology, Laboratory for Neuroimmunology, KU Leuven, Leuven, Belgium.
16. Department of Neurology, Danish Multiple Sclerosis Center, Copenhagen University Hospital, Copenhagen, Denmark.
17. Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland.
18. Département de Neurologie, INSERM UMR S 975 CRICM, UPMC, Pitié-Salpêtrière, Paris, France.
19. Department of Neurology, Klinikum rechts der Isar, Technische Universität München, Munich, Germany.
20. Munich Cluster for Systems Neurology (SyNergy), Munich, Germany.
21. German Competence Network Multiple Sclerosis (KKNMS), Munich, Germany.
22. Department of Clinical Sciences, Danderyd Hospital, Karolinska Institutet, Stockholm, Sweden.
23. Focus Program Translational Neuroscience (FTN), Rhine Main Neuroscience Network (rmn2), Johannes Gutenberg University-Medical Center, Mainz, Germany.
24. Department of Health Sciences and Interdisciplinary Research Center of Autoimmune Diseases (IRCAD), University of Eastern Piedmont, Novara, Italy.
25. Laboratory of Genetics of Neurological complex disorders, Institute of Experimental Neurology (INSPE), Division of Neuroscience, San Raffaele Scientific Institute, Milan, Italy.
26. Department of Neurology, Institute of Experimental Neurology (INSPE), Division of Neuroscience, San Raffaele Scientific Institute, Milan, Italy.
27. Menzies Research Institute Tasmania, University of Tasmania, Tasmania, Australia.
28. Department of Neurology, Oslo University Hospital, Ullevål, Oslo, Norway.
29. University of Oslo, Oslo, Norway.
30. Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden.

31. Department of Neurology, University of California at San Francisco, Sandler Neurosciences Center, San Francisco, California, USA.
32. Department of Neurology, MS Center ErasMS, Erasmus MC, Rotterdam, The Netherlands.
33. Division of Epidemiology, Genetic Epidemiology and Genomics Laboratory, School of Public Health, University of California, Berkeley, California, USA.
34. California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, California, USA.
35. Kaiser Permanente Division of Research, Oakland, California, USA.
36. Lists of authors and members appear in the Supplementary Note.
37. Laboratory of Molecular Medicine and Biotechnology, Don C. Gnocchi Foundation ONLUS, IRCCS S. Maria Nascente, Milan, Italy.
38. Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.
39. Université Rennes 1, Rennes, France.
40. Medical Research Council Biostatistics Unit, Cambridge, UK.
41. Department of Transfusion Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.
42. School of Medicine, Griffith University, Gold Coast, Queensland, Australia.
43. CHU Pellegrin, Université Bordeaux 2, Bordeaux, France.
44. University of Melbourne, Victoria, Melbourne, Australia.
45. Department of Neurology, Box Hill Hospital, Monash University, Victoria, Australia.
46. Service de Neurologie, CHRU Montpellier, Montpellier, France.
47. Plateforme Post-Génomique P3S, UPMC-INSERM, Paris, France.
48. Department of Neuroscience, MS Center, Azienda ospedaliera Città della Salute e della Scienza di Torino, Turin, Italy.
49. University of Turin, Turin, Italy.
50. Service de Neurologie, Hôpital Avicenne, Bobigny, France.
51. Department of Health Sciences, San Paolo Hospital and Filarete Foundation, University of Milan, Milan, Italy.
52. Service de Neurologie, CHU de Caen and INSERM U 919-GIP Cyceron, Caen, France.
53. Department of Neurology, Multiple Sclerosis Division, Miller School of Medicine, University of Miami, Miami, Florida, USA.
54. Neurologia 2 - CRESM, AOU San Luigi, Orbassano, Turin, Italy.
55. Department of Neurology, Royal Victoria Infirmary, Newcastle upon Tyne, UK.
56. Department of Neurology, University of Tampere, Medical School, Tampere, Finland.
57. Division of Clinical Neurology, Nottingham University Hospital, Nottingham, UK.
58. Florey Institute of Neuroscience and Mental Health, University of Melbourne, Victoria, Australia.
59. Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany.
60. Department of Pathophysiology and Transplantation, Neurology Unit, University of Milan, Milan, Italy.
61. Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy.
62. KORAgen, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Institute of Genetic Epidemiology, Neuherberg, Germany.
63. Department of Clinical Neurology, The Ipswich Hospital NHS Trust, Ipswich, UK.
64. Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA.
65. Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
66. Wellcome Trust Clinical Research Facility, Southampton General Hospital, Southampton, UK.
67. Department of Neurology, Aberdeen Royal Infirmary, Aberdeen, UK.
68. Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden.

69. Department of Medicine at Karolinska University Hospital Solna, Atherosclerosis Research Unit, Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden.
70. Department of Neurology, Hull Royal Infirmary, Hull, UK.
71. Department of Neurology, Royal Devon and Exeter Foundaton Trust Hospital, Exeter, Devon, UK.
72. Keele University Medical School, University Hospital of North Staffordshire, Stoke-on-Trent, UK.
73. UCL Genetics Institute (UGI), University College London, London, UK.
74. Department of Neurology, Poole General Hospital, Poole, UK.
75. Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK.
76. Institute for Neuroimmunology and Clinical MS Research (inims), Center for Molecular Neurobiology (ZMNH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany.
77. Department of Neuroimmunology and MS Research, Neurology Clinic, University Hospital Zürich, Zürich, Switzerland.
78. Department of Neurology, Division of Clinical Neurology, Leicester Royal Infirmary, Leicester, UK.
79. Australian Neuromuscular Research Institute, University of Western Australia, Western Australia, Australia.
80. Murdoch University, Western Australia, Australia.
81. Melbourne Neuroscience Institute, University of Melbourne, Victoria, Australia.
82. Department of Neurology, Seinäjoki Central Hospital, Seinäjoki, Finland.
83. IFM Bioinformatics, Linköping University, Linköping, Sweden.
84. Fondation Jean Dausset - Centre d'Etude du Polymorphisme Humain, Paris, France.
85. Commissariat à l'Energie Atomique, Institut Genomique, Centre National de Génotypage, Evry, France.
86. McGill University and Genome Quebec Innovation Center, Montreal, Canada.
87. Service de Neurologie, Hôpital Pasteur, CHRU Nice, France.
88. Hunter Medical Research Institute, University of Newcastle, New South Wales, Australia.
89. Department of Neurology, Ospedale Maggiore, Novara, Italy.
90. Department of Medical Genetics, Oslo University Hospital, Ullevål, Oslo, Norway.
91. Department of Vertebrate Genomics, Neuropsychiatric Genetics Group, Max Planck Institute for Molecular Genetics, Berlin, Germany.
92. Department of Clinical Neurosciences and Rehabilitation, Institute of Neuroscience and Physiology, Sahlgrenska Academy, Göteborgs Universitet, Göteborg, Sweden.
93. Department of Psychiatry and Human Behavior, School of Medicine, University of California, Irvine, California, USA.
94. Department of Pharmacological and Biomolecular Sciences, University of Milan, Milan, Italy.
95. Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland.
96. Canterbury District Health Board, Christchurch, New Zealand.
97. Department of Neurology, Queen Elizabeth Medical Centre, Edgbaston, Birmingham, UK.
98. Department of Neurology, The Norwegian Multiple Sclerosis Registry and Biobank, Haukeland University Hospital, Bergen, Norway.
99. National Multiple Sclerosis Center Melsbroek, Melsbroek, Belgium.
100. Neurology Department, Charing Cross Hospital, London, UK.
101. Department of Clinical Sciences, Lund University, Lund, Sweden.
102. Department of Neurology, Royal Hallamshire Hospital, Sheffield, UK.
103. Department of Neurology, University of Oulu, Oulu, Finland.
104. Department of Neurology, University Hospital of Oulu, Oulu, Finland.
105. Laboratory of Neurobiology, Vesalius Research Center, Leuven, Belgium.
106. Experimental Neurology, Leuven Research Institute for Neurodegenerative Diseases (LIND), University of Leuven (KU Leuven), Leuven, Belgium.
107. Department of Neurology, University Hospitals Leuven, Leuven, Belgium.

108. Institute of Psychological Medicine and Clinical Neuroscience, Cardiff University, University Hospital of Wales, Cardiff, UK.
109. Department of Neurology, Greater Manchester Neurosciences Centre, Salford Royal NHS Foundation Trust, Salford, UK.
110. Department of Neuroscience, Centre for Experimental Neurological Therapies, Mental Health and Sensory Organs, Sapienza Università di Roma, Rome, Italy.
111. Department of Neurology, Royal Preston Hospital, Preston, UK.
112. Department of Neurology, Norfolk and Norwich Hospital, Norwich, UK.
113. Department of Neurology, Flinders University, Adelaide, South Australia, Australia.
114. Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway.
115. Department of Neurology, Peterborough City Hospital, Peterborough, UK.
116. Neurology and Stroke Unit, San Francesco Hospital, Nuoro, Italy.
117. Department of Neurology, Helsinki University Central Hospital and Molecular Neurology Programme, Biomedicum, University of Helsinki, Helsinki, Finland.
118. Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands.
119. Division of Applied Health Sciences, University of Aberdeen, Foresterhill, Aberdeen, UK.
120. Institute of Clinical Neurosciences, University of Bristol, Frenchay Hospital, Bristol, UK.
121. Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK.
122. Institut für Humangenetik, Technische Universität München, Munich, Germany.
123. Institut für Humangenetik, Helmholtz Zentrum München, Munich, Germany.
124. Harvard NeuroDiscovery Center, Harvard Medical School, Boston, Massachusetts, USA.
125. Department of Immunobiology, Yale University School of Medicine, New Haven, Connecticut, USA.
126. These authors contributed equally to this work.

Corresponding author:

Jacob L. McCauley

jmccauley@med.miami.edu

Author Contributions

M.F.D., D. Booth, A.O., J.S., B. Fontaine, B.H., C. Martin, F.Z., S.D.'A., F.M.-B., B.T., H.F.H., I. Kockum, J. Hillert, T.O., J.R.O., R.H., L.F.B., C. Agliardi, L.A., L. Bernardinelli, V.B., S.B., B.B., L. Brundin, D. Buck, H. Butzkeuven, W. Camu, P.C., E.G.C., I.C., G.C., I.C.-R., B.A.C.C., G.D., S.R.D., A.D.S., B.D., M.D., I.E., F.E., N.E., J.F., A.F., I.Y.F., D.G., C. Graetz, A. Graham, C. Guaschino, C. Halfpenny, G. Hall, J. Harley, T.H., C. Hawkins, C. Hillier, J. Hobart, M.H., I.J., A.J., B.K., A. Kermode, T. Kilpatrick, K.K., T. Korn, H.K., C.L.-F., J.L.-S, M.H.L., M.A.L., G.L., B.A.L., C.M.L., F.L., J. Lycke, S.M., C.P.M., R.M., V.M., D.M., G. Mazibrada, J.M., K.M., G.N., R.N., P.N., F.P., S.E.P., H.Q., M. Reunanen, W.R., N.P.R., M. Rodegher, D.R., M. Salvetti, F.S., R.C.S., C. Schaefer, S. Shaunak, L.S., S. Shields, V.S., M. Slee, P.S.S., M. Sospedra, A. Spurkland, V.T., J.T., A.T., P.T., C.V.D., E.M.V., S.V., J.S.W., A.W., J.F.W., J.Z., E.Z., J.L.H., M.A.P.-V., G.S., D.H., S.L.H., A.C., P.D.J., S.J.S. and J.L.M. were involved with case ascertainment and phenotyping . A. Kemppinen, D. Booth, A. Goris, A.O., B. Fontaine, S.D.'A., F.M.-B., H.F.H., I. Kockum, M.B., J.R.O., L.F.B., IIBDGC, H.B.S., A. Baker, N.B., L. Bergamaschi, I.L.B., P.B., D. Buck, S.J.C., L. Corrado, L. Cosemans, I.C.-R., V.D., J.F., A.F., V.G., I.J., I. Konidari, V.L., C.M.L., M. Lindén, J. Link, C. McCabe, I.M., H.Q., M. Sorosina, E.S., H.W., P.D.J., S.J.S. and J.L.M. processed the DNA. A. Kemppinen, A.O., B. Fontaine, M.B., R.H., L.F.B., WTCCC2, IIBDGC, R.A., H.B.S., N.B., T.M.C.B., H. Blackburn, P.B., W. Carpentier, L. Corrado, I.C.-R., D.C., V.D., P. Deloukas, S.E., A.F., H.H., P.H., A. Hamsten, S.E.H., I.J., I. Konidari, C.L., M. Larsson, M. Lathrop, F.M., I.M., J.M., H.Q., F.S., M. Sorosina, C.V.D., J.W., D.H., P.D.J., S.J.S. and J.L.M. conducted and supervised the genotyping of samples. A.H.B., N.A.P., D.K.X., M.F.D., A. Kemppinen, C.C., T.S.S., C. Spencer, M.B., IIBDGC, C. Anderson, S.E.B., A.T.D., P. Donnelly, B. Fiddes, P.G., G. Hellenthal, S.E.H., L.M., M.P., N.C.S.-B., J.L.H., M.A.P.-V., G. McVean, P.D.J., S.J.S. and J.L.M. performed the statistical analysis . A.H.B., N.A.P., D.K.X., M.F.D., A. Kemppinen, C.C., T.S.S., C. Spencer, D. Booth, A. Goris, A.O., J.S., B. Fontaine, B.H., F.Z., S.D.'A., F.M.-B., H.F.H., I. Kockum, M.B., R.H., L.F.B., C. Agliardi, M.A., C. Anderson, R.A., H.B.S., A. Baker, G.B., N.B., J.B., C.B., L. Bernardinelli, A. Berthele, V.B., T.M.C.B., H. Blackburn, I.L.B., B.B., D. Buck, S.J.C., W. Camu, P.C., E.G.C., I.C., G.C., L. Corrado, L. Cosemans, I.C.-R., B.A.C.C., D.C., G.D., S.R.D., P. Deloukas, A.D.S., A.T.D., P. Donnelly, B.D., M.D., S.E., F.E., N.E., B. Fiddes, J.F., A.F., C.F., D.G., C. Gieger, C. Graetz, A. Graham, V.G., C. Guaschino, A. Hadjixenofontos, H.H., C. Halfpenny, P.H., G. Hall, A. Hamsten, J. Harley, T.H., C. Hawkins, G. Hellenthal, C. Hillier, J. Hobart, M.H., S.E.H., I.J., A.J., B.K., I. Konidari, H.K., C.L., M. Larsson, M. Lathrop, C.L.-F., M.A.L., V.L., G.L., B.A.L., C.M.L., F.M., C.P.M., R.M., V.M., G. Mazibrada, C. McCabe, I.M., L.M., K.M., R.N., M.P., S.E.P., H.Q., N.P.R., M. Rodegher, D.R., M. Salvetti, N.C.S.-B., R.C.S., C. Schaefer, S. Shaunak, L.S., S. Shields, M. Sospedra, A. Strange, J.T., A.T., E.M.V., A.W., J.F.W., J.W., J.Z., J.L.H., A.J.I., G. McVean, P.D.J., S.J.S. and J.L.M. collected and managed the project data. A.H.B., N.A.P., M.F.D., A. Kemppinen, C.C., T.S.S., C. Spencer, J.S., B.H., F.Z., S.D.'A., F.M.-B., H.F.H., J. Hillert, T.O., M.B., J.R.O., R.H., L.F.B., L.A., C. Anderson, G.B., J.B., C.B., A. Berthele, E.G.C., G.C., P. Donnelly, F.E., C.F., C. Gieger, C. Graetz, G. Hellenthal, M.J., T. Korn, M.A.L., R.M., M.P., M. Sospedra, A. Spurkland, A. Strange, J.W., J.L.H., M.A.P.-V., A.J.I., G.S., D.H., S.L.H., A.C., G. McVean, P.D.J., S.J.S. and J.L.M. contributed to the study concept and design. A.H.B., N.A.P., D.K.X., G. McVean, P.D.J., S.J.S. and J.L.M. prepared the manuscript. All authors reviewed the final manuscript.

The authors have no competing financial interests

Online Methods

ImmunoChip data (discovery set).

Details of case ascertainment, processing and genotyping for the discovery phase are provided in the **Supplementary Note** (**Supplementary Table 9**). Genotype calling for all samples was performed using Opticall.[40] Samples that performed poorly or were determined to be related were removed (**Supplementary Table 10**). The data were organized in 11 country level strata: ANZ (Australia + New Zealand), Belgium, Denmark, Finland, France, Germany, Italy, Norway, Sweden, United Kingdom (UK), and the United States of America (USA). SNP level quality control (**Supplementary Table 11**) and population outlier identification using principal components analysis (**Supplementary Fig. 97**) were done in each stratum separately.

Discovery set analysis.

We applied logistic regression, assuming a per-allelic genetic model per data set, including the first five principal components as covariates to correct for population stratification (**Supplementary Table 12** lists the per data set genomic inflation factors, λ). We then performed an inverse-variance meta-analysis of the 11 strata, under a fixed effects model, as implemented in PLINK.[41] To be more conservative and account for any residual inflation in the test statistic, we applied the genomic control equivalent to the per-SNP standard error in each stratum. Specifically, we corrected the SNP standard errors by multiplying them with the square root of the raw genomic inflation factor λ, per data set, if the λ was >1.

Within the designated fine-mapping intervals, we applied a forward stepwise logistic regression to identify statistically independent effects. The primary SNP in each interval was included as a covariate, and the association analysis was repeated for the remaining SNPs. This process was repeated until no SNPs reached the minimum level of significance (p-value $<1 \times 10^{-4}$). Outside of the designated fine-mapping intervals, all SNPs having a p-value $<1 \times 10^{-4}$ were identified and grouped into sets based on a physical distance of less than 2Mb and a similar stepwise regression model was applied. Any SNPs to enter the model with p-value $<1 \times 10^{-4}$ after conditioning were considered statistically independent primary signals.

In addition, because of the close physical proximity between some fine-mapping intervals and SNP sets, independence was tested for all identified signals within 2Mb of one another. The and cluster plots (**Supplementary Fig. 98**) of all independent SNPs were examined, and the SNP was excluded if unsatisfactory. If any SNP was excluded, the forward stepwise logistic regression within that fine-mapping

interval or SNP set was repeated after removal of the SNP. During this process, 17 additional SNPs were excluded based on cluster or forest plot review.

Replication Set.

The replication phase included GWAS data organized into 15 strata. Within each stratum, poorly performing samples (call rate < 95%, gender discordance, excess heterozygosity) and poorly performing SNPs (Hardy-Weinberg equilibrium (HWE) p-value $<1 \times 10^{-6}$, minor allele frequency (MAF) < 1%, call rate < 95%) were removed. Principal components analysis was performed to identify population outliers per stratum, and the genomic control inflation factor was < 1.1 for each. The data included in the final discovery and replication analyses are summarized in **Supplementary Table 13** and **Supplementary Table 14**. All the samples used in the replication set were unrelated to those in the discovery set; verified by identity-by-descent analysis.

We attempted replication of all non-MHC independent signals that reached a discovery p-value of $<1 \times 10^{-4}$ in a meta-analysis set of GWAS. Each data set was imputed to the 1000 Genomes European phase I (a) panel using BEAGLE[42] to maximize the overlap between the Immunochip SNP content and the GWAS data. Post-imputation genotypic probabilities were used in a logistic regression model, per stratum, to estimate SNP effect sizes and p-values. By using the post-imputation genotypic probabilities, we penalized SNPs that didn't have good imputation quality, thus ensuring a conservative analysis. Furthermore, we accounted for population stratification in each data set by including the first five principal components in the logistic model. We then meta-analysed the effect size and respective standard errors of the 15 strata using a fixed effects model inverse-variance method. We applied the genomic control equivalent to the per-SNP standard error in each stratum, controlling for the respective genomic inflation factor λ (**Supplementary Table 14**).

To replicate the primary SNPs per identified signal in the discovery phase, we used the replication effect size and respective standard error. For the secondary and tertiary SNPs, we fitted the same exact models as in the discovery phase, per data set. We then performed fixed effects meta-analysis to estimate an effect size that corresponds to the same logistic model. In the case that a SNP was not present in the replication set, we replaced it with a perfectly tagging SNP, i.e. a SNP that had r^2 and D' equal to 1. If a perfectly tagging SNP was not available, we selected a SNP that had equivalent MAF and the highest possible r^2 and D'. Estimation of r^2 and D' for this objective were based on the ImmunoChip control samples.

Joint analysis (discovery and replication sets).

The discovery and replication phase effect sizes and respective standards errors were meta-analysed under a fixed effects model. A SNP was considered replicated when all three of the following criteria were met: 1)

replication p-value <5.0 x 10^{-2}, 2) joint p-value <5 x 10^{-8}, and 3) the joint p-value was more statistically significant than the discovery p-value. SNPs that reached a p-value of <1 x 10^{-6} but did not pass the genome-wide threshold, were coined suggested if the above criteria 1) and 3) were met.

Fine-mapping of association signals.

To fine-map signals of association we used a combination of imputation and Bayesian methodology.[23] Around each of the 97 associated SNPs, 2Mb were isolated in the discovery and replication phase UK data as well as the European samples from the Phase 1 1000G.[28] Forming the single largest cohort, only UK samples were considered to minimize the effects of differential imputation quality between populations of different ancestry. In addition to the previous quality control, SNPs with failed alignment or a difference in MAF >10% between the typed cohorts and the 1000G samples, MAF < 1%, or HWE p-value <1.0 x 10^{-4} were removed.

Imputation was performed separately for the UK discovery and replication cohorts on each 2Mb region using the default settings of IMPUTEv2.[43,44] Missing genotypes in the genotyped SNPs were not imputed, and any imputed SNP that failed the HWE and MAF threshold was subsequently removed. We carried out frequentist and Bayesian association tests on all SNPs in each cohort separately, assuming additivity, using the default settings of SNPTESTv2.[45] Frequentist fixed-effect meta-analysis was carried out using the software META.[46] Bayesian meta-analysis was carried out using an independence prior (near-identical results were obtained using a fixed-effect Bayesian meta-analysis).

To identify regions where reliable fine-mapping could be achieved, we used the information score (INFO, obtained from IMPUTEv2) as identified from the 1000G samples. Specifically, we measured the fraction of variants with both $r^2 > 0.5$ and $r^2 > 0.8$ to the primary associated variant, having greater than 50% and 80% INFO scores respectively. Regions where any SNP with $r^2 > 0.5$ had INFO < 50% were excluded. We also excluded regions where the top hit from imputation had an INFO score less than 80%. Regions were considered to be fine-mapped with high quality when all variants with $r^2 > 0.8$ had at least 80% INFO. Within these regions, we excluded variants where the inferred direction of association was opposite in the UK discovery and replication cohorts.

To measure the posterior probability that any single variant drives association, we calculated the Bayes Factor. Under the assumption that there is a single causal variant in the region, this is proportional to the probability that the variant drives the association.[23] We identified the smallest set of variants that contained 90% and 50% of the posterior probability. We called a region successfully and consistently fine-mapped if there were at most five variants in the 50% confidence interval and the top SNP from the frequentist analysis lived in the 90% confidence interval. For these regions, we annotated variants with information about evolutionary conservation, predicted coding consequence, regulation, published associations to expression

or DNase I hypersensitive sites using ANNOVAR,[47] VEP,[24] and the eQTL browser, a recent immune cell expression study[20], and other literature.

Gene Ontology.

To determine the GO processes for which our associated variants were involved, we used MetaCore from Thomson Reuters. We annotated the processes for the unique genes within 50Kb of the variants.

Cross disease comparison.

In order to explore the potential overlap with variants identified across other autoimmune diseases, we calculated the percentage overlap of reported variants found in other ImmunoChip reports to our ImmunoChip results. The top variants reported as either novel or previously known in other ImmunoChip reports were compared with the 110 variants representing both our novel and previous discoveries in multiple sclerosis. In order for a signal to be considered as overlapping, we required an $r^2 \geq 0.8$ using the Pairwise LD function of the SNAP tool in European samples.[31]

Secondary analyses.

We performed a severity based analysis of MSSS in cases only from the discovery phase (**Supplementary Fig. 99**). In addition, a transmission disequilibrium test was done in 633 trios to test for transmission of the 97 identified risk alleles (**Supplementary Fig. 100**). Details are given in the **Supplementary Note**.

References

1. Gourraud, P.A., Harbo, H.F., Hauser, S.L. & Baranzini, S.E. The genetics of multiple sclerosis: an up-to-date review. *Immunol Rev* **248**, 87-103 (2012).
2. Nylander, A. & Hafler, D.A. Multiple sclerosis. *J Clin Invest* **122**, 1180-8 (2012).
3. Compston, A. et al. *McAlpine's Multiple Sclerosis*, (Churchill Livingstone, London, 2006).
4. Dyment, D.A., Yee, I.M., Ebers, G.C. & Sadovnick, A.D. Multiple sclerosis in stepsiblings: recurrence risk and ascertainment. *J Neurol Neurosurg Psychiatry* **77**, 258-9 (2006).
5. Hemminki, K., Li, X., Sundquist, J., Hillert, J. & Sundquist, K. Risk for multiple sclerosis in relatives and spouses of patients diagnosed with autoimmune and related conditions. *Neurogenetics* **10**, 5-11 (2009).
6. Jersild, C., Svejgaard, A. & Fog, T. HL-A antigens and multiple sclerosis. *Lancet* **1**, 1240-1 (1972).
7. IMSGC. Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study. *N Engl J Med* **357**, 851-62 (2007).
8. De Jager, P.L. et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* **41**, 776-82 (2009).
9. IMSGC & WTCCC2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214-9 (2011).
10. Patsopoulos, N.A. et al. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* **70**, 897-912 (2011).
11. IMSGC. Evidence for polygenic susceptibility to multiple sclerosis--the shape of things to come. *Am J Hum Genet* **86**, 621-5 (2010).
12. Baranzini, S.E. The genetics of autoimmune diseases: a networked perspective. *Curr Opin Immunol* **21**, 596-605 (2009).
13. Cotsapas, C. et al. Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet* **7**, e1002254 (2011).
14. Cortes, A. & Brown, M.A. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* **13**, 101 (2011).
15. Jostins, L. et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119-24 (2012).
16. Krzywinski, M. et al. Circos: an information aesthetic for comparative genomics. *Genome Res* **19**, 1639-45 (2009).
17. Willis, T.G. et al. Bcl10 is involved in t(1;14)(p22;q32) of MALT B cell lymphoma and mutated in multiple tumor types. *Cell* **96**, 35-45 (1999).
18. Yan, J. & Greer, J.M. NF-kappa B, a potential therapeutic target for the treatment of multiple sclerosis. *CNS Neurol Disord Drug Targets* **7**, 536-57 (2008).
19. Wegener, E. & Krappmann, D. CARD-Bcl10-Malt1 signalosomes: missing link to NF-kappaB. *Sci STKE* **2007**, pe21 (2007).
20. Fairfax, B.P. et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet* **44**, 502-10 (2012).
21. Lill, C.M. et al. Genome-wide significant association of ANKRD55 rs6859219 and multiple sclerosis risk. *J Med Genet* **50**, 140-3 (2013).
22. Maier, L.M. et al. IL2RA genetic heterogeneity in multiple sclerosis and type 1 diabetes susceptibility and soluble interleukin-2 receptor production. *PLoS Genet* **5**, e1000322 (2009).
23. Maller, J.B. et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat Genet* **44**, 1294-301 (2012).
24. McLaren, W. et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**, 2069-70 (2010).

25. Gregory, A.P. et al. TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature* **488**, 508-11 (2012).
26. De Jager, P.L. et al. The role of the CD58 locus in multiple sclerosis. *Proc Natl Acad Sci U S A* **106**, 5264-9 (2009).
27. Malmestrom, C. et al. Serum levels of LIGHT in MS. *Mult Scler* (2012).
28. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
29. Schaub, M.A., Boyle, A.P., Kundaje, A., Batzoglou, S. & Snyder, M. Linking disease associations with regulatory information in the human genome. *Genome Res* **22**, 1748-59 (2012).
30. Davydov, E.V. et al. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol* **6**, e1001025 (2010).
31. Johnson, A.D. et al. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**, 2938-9 (2008).
32. Juran, B.D. et al. Immunochip analyses identify a novel risk locus for primary biliary cirrhosis at 13q14, multiple independent associations at four established risk loci and epistasis between 1p31 and 7q32 risk variants. *Hum Mol Genet* **21**, 5209-21 (2012).
33. Liu, J.Z. et al. Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nat Genet* **44**, 1137-41 (2012).
34. Trynka, G. et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet* **43**, 1193-201 (2011).
35. Eyre, S. et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat Genet* **44**, 1336-40 (2012).
36. Tsoi, L.C. et al. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nat Genet* **44**, 1341-8 (2012).
37. Cooper, J.D. et al. Seven newly identified loci for autoimmune thyroid disease. *Hum Mol Genet* **21**, 5202-8 (2012).
38. Ban, M. et al. Replication analysis identifies TYK2 as a multiple sclerosis susceptibility factor. *Eur J Hum Genet* **17**, 1309-13 (2009).
39. Ban, M. et al. A non-synonymous SNP within membrane metalloendopeptidase-like 1 (MMEL1) is associated with multiple sclerosis. *Genes Immun* **11**, 660-4 (2010).
40. Shah, T.S. et al. optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. *Bioinformatics* **28**, 1598-603 (2012).
41. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
42. Browning, B.L. & Yu, Z. Simultaneous genotype calling and haplotype phasing improves genotype accuracy and reduces false-positive associations for genome-wide association studies. *Am J Hum Genet* **85**, 847-61 (2009).
43. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).
44. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* **1**, 457-70 (2011).
45. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat Rev Genet* **11**, 499-511 (2010).
46. Liu, J.Z. et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet* **42**, 436-40 (2010).
47. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).